A Magnetotelluric Investigation of Hawaiian Groundwater Systems in Kona, Hawai‘i Island, Hawai‘i, including Uncertainty Quantification

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE
IN
GEOLOGY AND GEOPHYSICS

December 2019

by
Taylor Viti

Thesis committee:
Niels Grobbe, Chairperson
Neil Frazer
Nicole Lautze

Keywords: Groundwater, magnetotelluric, Archie's law, uncertainty quantification

Acknowledgments

I would like to thank Thomas Martin, Ted Brennis, Young-Ho Seo, Dr. Ahmed Elshall, Dr. Niels Grobbe, and Dr. Donald Thomas, for their help in the field collecting the magnetotelluric data described in this work. The data collection effort also would not have been possible without Britt Craven, of the Palani Ranch Co., whom I'd like to thank for allowing us to conduct this survey on his land (with the trust to allow us to do so unsupervised).

I would also like to thank my advisor, Dr. Niels Grobbe, for his tutelage over the past two years, for giving me the opportunity to participate in multiple geophysical surveys, and for always allowing me to have a voice in our discussions. I am additionally indebted to my committee members Dr. Neil Frazer, and Dr. Nicole Lautze, for their guidance throughout my master's degree.

Lastly, I thank my mother, Rubylyn Mata-Viti, for always giving me her love, support, and patience, and my significant other Diamond Tachera, for being a continuous source of strength and inspiration. Throughout my life, I have somehow always found myself in the company of strong, motivated, and enduring women, with Diamond and my mother being no exceptions to this rule.

This study was funded by the NSF EPSCoR ʻIke Wai project. Support for the Hawaiʻi EPSCoR Program is provided by the National Science Foundation's Research Infrastructure Improvement (RII) Track-1: ʻIke Wai: Securing Hawaiʻi's Water Future Award #OIA-1557349.

Abstract

We investigate the spatial distribution of groundwater at high elevations in Hawai'i Island's Keauhou aquifer system, where a sparse distribution of water wells indicates anomalously high head levels that are likely caused by a geological structure of unknown nature. To study this system, we have conducted a magnetotelluric (MT) survey on a transect climbing the southwestern flank of the Hualālai shield volcano. We carry-out a 2D deterministic inversion using MARE2DEM (a freely available MT modeling program), producing an image of the bulk electrical resistivity distribution in the subsurface. We explore the use of 1D, station-by-station, Bayesian inversions (using our own software), in an attempt to supplement the limited amount of water wells at higher elevations, by creating 'virtual water wells' from MT stations that can be used to infer e.g. head-level information. Even though we can infer information about the subsurface groundwater distribution with reasonable accuracy, we infer that the vertical resolution of the MT data is most likely insufficient to be used as actual virtual water well head-level data. However, the MT results are anticipated to be useful for roughly validating and constraining hydrological models in terms of e.g. different hydrogeological regimes.

Comparing the 1D inversion results with the 2D results, we find that the 1D inversion

results match the 2D results reasonably well for most stations, especially until the depth at which seawater associated resistivities are encountered. Using the uncertainty quantification of the Bayesian inversions, we determine that depths of up to approximately 1 km beneath mean sea level are well constrained by the data. To first order, the 2D image exhibits a layered resistivity structure. Based on literature values for the resistivities of Hawaiian basalts under various saturation conditions, these layers are consistent with (in order of increasing depth), unsaturated Hawaiian basalt, freshwater-saturated Hawaiian basalt, and seawater-saturated Hawaiian basalt. A large head gradient delineating inland high-level heads from coastal low-level heads is also evident in the image as an apparent discontinuity in the freshwater valued resistivities. Resistivities associated with seawater-saturated basalts occur in the high-level aquifer starting at a depth of approximately 1 km beneath mean sea level, suggesting that the high-level aquifer is a vertically extensive body of fresh water that is hydraulically continuous with seawater, most likely via a brackish water mixing zone.

In light of the virtual water-well approach, the resistivities from the 2D image are inverted for porosity and saturant salinity, using an Archie's law-based Bayesian inversion routine that we have developed. The resulting salinity profile reflects an increase in salinity with depth, but with very few samples meeting or exceeding the salinity of seawater. Due to the non-identifiability of the Archie's law parameters, incorporating sufficient, high-quality prior information is probably key when trying to obtain useful outcomes, for example by constraining parameter ranges in terms of depth range, using prior porosity information from core samples, or using prior salinity information from water wells in the region.

Table of Contents

List of Figures

Acronyms

AMSL above mean sea level

BMSL below mean sea level

CI central interval

EM electromagnetic

ERT electrical resistivity tomography

GPS global positioning system

MCMC Markov chain Monte-Carlo

MSL mean sea level

MT magnetotelluric

PPT parts per thousand

TDEM time-domain electromagnetics

UH University of Hawai‘i at Mānoa

UQ uncertainty quantification

Chapter 1

Introduction

1-1 Thesis overview

The Hawaiian islands depend predominantly on local aquifers for drinkable water (Gingerich and Oki, 2000) (Lau and Mink, 2006), underscoring the importance of understanding how these resources behave. In contrast, Izuka et al. (2018) described high-elevation groundwater in the Big Island of Hawai'i's Keauhou and Kīholo aquifers, as a 'groundwater enigma' (Izuka et al., 2018, Page 105). Large hydraulic gradients are observed in these aquifers, across a line termed the high-low divide (or basal divide), trending roughly parallel to the coastline (Fig. 1-1). Water levels in wells drilled on its coastal side are on the order of a few meters above mean sea level (AMSL), while a few kilometers inland they instead range from tens to hundreds of meters AMSL. This large head gradient suggests that geological structures in the subsurface, whose origins are largely unknown, form a hydrological barrier to groundwater flow towards the coast. Proposed geological mechanisms for the divide have included complexes of nearly impermeable dikes, sequences of 'massive' low permeability lava flows, and a fault system trending parallel to the coast (Oki, 1999). Evidence suggesting that the high-level resource is comprised of interconnected compartments was presented by Bauer (2003), with the author concluding that multiple geological mechanisms are likely to be at play. More recently, Fackrell (2016, Fig. 3.7c) presented a model with shallow perched groundwater resting hundreds of meters above a vertically extensive confined aquifer. Although the high-level classification has been used in the past

to describe freshwater that is not in contact with seawater (Mink and Lau, 1993) (Bauer, 2003) (Lau and Mink, 2006), many of these models involve interactions between the high-level resource and the seawater presumed to underlie the coastal resource. Water wells in the high-level aquifer are clustered close to the divide line, making direct observations of the subsurface, with which hydrological models of these conceptual models could be validated and/or constrained, virtually non-existent.

Because the bulk electrical resistivity of a rock changes when it is saturated with water (and also when the resistivity of the water changes), electromagnetic (EM) geophysical methods such as electrical resistivity tomography (ERT), time-domain electromagnetics (TDEM), and the magnetotelluric method (MT), which produce estimates of the subsurface distribution of bulk electrical resistivity, can provide particularly valuable information about high level groundwater. Rai and Manghnani (1981) tested the bulk electrical resistivity of Hawaiian basalt samples of varying porosity, and saturated with four solutions of varying salinity, demonstrating a correlation between bulk resistivity, porosity, salinity, and confining pressure. Pierce and Thomas (2009) conducted an MT groundwater exploration survey in the Big Island's Humu'ula Saddle region, reporting resistivities for dry- and freshwater-saturated basalts based on borehole water-level observations. Other factors potentially influencing bulk resistivity include temperature, chemical alteration, the presence of clay minerals, and the rock's surface conductance (Revil et al., 2016)(Revil et al., 2018), which we do not consider in this work.

Utilizing results from these studies as a basis for interpreting resistivity images in volcanic environments, we can use the distribution of bulk resistivity in the high-level aquifer as a proxy for the presence and salinity of water in the subsurface, and in turn infer if the high-level freshwater body is truncated with seawater, possibly via a brackish mixing zone. Presuming that the high-level resource is a vertically extensive body of fresh groundwater, juxtaposed with a thin basal coastal aquifer, we should additionally be able to see the hydraulic discontinuity between the two regimes as a large change in resistivity across the divide line. Further inland, we may be able to identify perched groundwater bodies, provided that they are large enough to produce a noticeable signature in an MT resistivity image. Surface elevation on the high-level side of the divide ranges above 500 m AMSL, placing our potential fresh and saltwater targets outside of the depths of investigation for near-surface EM techniques like ERT and TDEM, but well within reach for an MT survey.

To image the electrical resistivity distribution in the Keauhou aquifer, we have conducted an MT survey along a transect trending perpendicular to the Kona coast (Fig. 1-1). We use

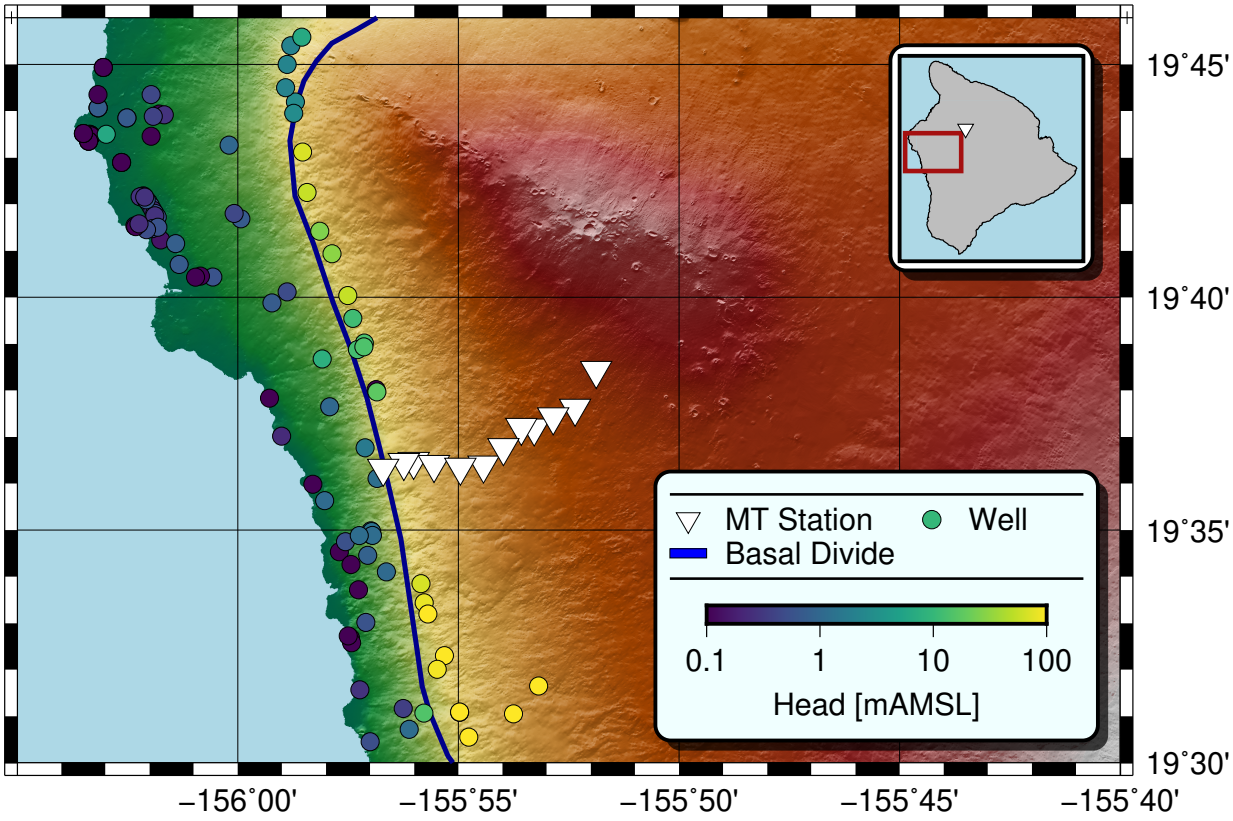


Figure 1-1: Map showing MT station deployments (white triangles), water wells and their initial head levels (colored circles), and the presumed location of the high-low divide (dark blue solid line) based on observed head levels (adapted from Fackrell (2016)). The remote reference's location is shown by a white triangle on the inset map. The red box on this inset map indicates the study site (main map). The contrast between high- and low-level well heads is strongest in the southern portion of the study site, where high-level heads are on the order of hundreds of meters AMSL.

the data from this survey to form a 2D image of the subsurface electrical resistivity beneath the transect. We also form 1D estimates of the posterior probability distribution of resistivity beneath each MT station location, which were used to determine the reliably interpretable regions of the 2D image. Using bulk resistivity values from Rai and Manghnani (1981) and Pierce and Thomas (2009) for dry, freshwater-saturated, and seawater-saturated Hawaiian basalts, we attempt to identify hydrogeological features in the image that are consistent with the vadose zone, freshwater-saturated rock, and seawater-saturated rock, in turn determining if seawater may be present in the high-level aquifer, along with the depths at which it may occur. Finally, in order to quantify salinity versus depth in the high-level aquifer, we develop a method for using resistivity data to estimate the posterior probability distributions of salinity and porosity in saturated formations. We apply this workflow to a subset of our 2D electrical resistivity image.

1-2 Site description

The Keauhou aquifer system is situated on the southwestern flank of the shield volcano Hualālai, and is comprised of Hualālai volcanics, with some evidence from wells outside Keauhou that at depth, the Hualālai flows interfinger with flows from Maunaloa, a younger shield volcano to the south (Bauer, 2003). The rocks in the subsurface are predominantly basalts, with core logs in the study area revealing some trachytes as well (Bauer, 2003, Table 2). In an archetypal basal Hawaiian groundwater system comprised of permeable flank lavas (a scenario presumed accurate for the low-level aquifer), the density contrast between fresh- and saltwater causes fresh groundwater to float on top of seawater. Assuming that the flow of the water is predominantly horizontal, the depth to the midpoint of the freshwater to saltwater transition zone can be approximated from head-level measurements using the Ghyben-Herzberg ratio (Lau and Mink, 2006, Appendix 6.1), which predicts a 40 to 1 ratio between the depth of the transition zone midpoint below mean sea level (MSL), and the head level above MSL, respectively. Substituting a maximum low-level head of 4.2 m AMSL (Bauer, 2003) into this relationship yields a maximum low-level aquifer thickness of 168 m. In contrast, applying the Ghyben-Herzberg ratio to the observed high-level heads yields high-level aquifer thicknesses on the order of kilometers.

Throughout the Hawaiian islands, high-level groundwater generally occurs as either dike-impounded water, in which low permeability, vertically-oriented intrusive formations restrict the lateral movement of freshwater, or as perched groundwater resting on low-

permeability layers such as ash deposits and unusually dense lava flows (Lau and Mink, 2006). Dike complexes are formed at or near a volcanoes rift zone, as magma migrating through the subsurface penetrates the country rock, with the resulting thin, sheet like structures oriented parallel to the rift zone. In our study site, the presumed location and orientation of the high-low divide are not associated with any of Hualālai's known rift zones. The presence of intrusives along the divide would thus require the existence of an undiscovered, buried rift zone (Bauer, 2003). We presume that intrusive formations, due to their low permeability, would tend to isolate the high-level aquifer from seawater intrusion.

Chapter 2

Magnetotellurics: theory and methods

2-1 The magnetotelluric method

The magnetotelluric method (MT) is a technique for characterizing the electrical resistivity distribution of Earth's subsurface. Natural variations in the Earth's magnetic field induce electrical currents in the subsurface (termed *telluric currents*), which in turn emit secondary electric and magnetic fields, such that surface measurements of the net electric and magnetic fields can be used to infer the subsurface spatial distribution of electrical resistivity (Chave and Jones, 2012). Data loggers equipped with magnetometers, ground dipole antennas, and GPS receivers, are deployed in the field to measure temporal variations in the Earth's electric and magnetic vector fields, which in the frequency domain are linearly related by the electromagnetic impedance tensor (Chave and Jones, 2012). Denoting the frequency-domain forms of the horizontal surface electric and magnetic vector fields as $\mathbf{E}$ and $\mathbf{H}$, respectively, and the impedance tensor as $\mathbf{Z}$, the relationship between these three quantities is given by

$$\begin{aligned} \mathbf{E} &= \mathbf{Z}\mathbf{H}, \\ \begin{pmatrix} E_x \\ E_y \end{pmatrix} &= \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix}, \end{aligned} \tag{2-1}$$

where the subscripts x and y occurring in the electric and magnetic vector field expressions denote orthogonal horizontal field components, with the x direction commonly aligned parallel to true North. In a 1D isotropic 'layer-cake' Earth model, the diagonal elements of $\mathbf{Z}$ are zero, and the off diagonals are the negatives of one another. In a 2D Earth model with isotropic resistivities, the diagonal elements are still both zero, and the off-diagonal elements are independent. In a 3D isotropic (or 2D anisotropic) Earth, all four components of $\mathbf{Z}$ are independent (Simpson and Bahr, 2005). The apparent resistivity ρ^a is a derived quantity representing the 'average' resistivity of the Earth at a given frequency. In a homogeneous halfspace, ρ^a is equal to the halfspace resistivity. Together with impedance phase ϕ, ρ^a is commonly used in place of $\mathbf{Z}$ for analysis, as ϕ and ρ^a are generally thought to be more intuitive than $\mathbf{Z}$, and have also been demonstrated to provide some computational benefit over $\mathbf{Z}$ (Wheelock et al., 2015). For each element Z_{ij} of $\mathbf{Z}$, apparent resistivity and phase are given by

$$\rho^a_{ij} = \frac{|Z_{ij}|^2}{\omega\mu_0}, \qquad \phi_{ij} = \mathrm{Arg}\{Z_{ij}\}\,. \tag{2-2}$$

Here, the subscripts ij denote index notation, where these subscripts can take the values 1, 2, and 3 corresponding to the three spatial directions in a right handed coordinate system (with positive x_3 oriented downwards). The vertical component of the magnetic vector field can additionally be related to the horizontal components by means of a transfer function known as the tipper (Chave and Jones, 2012). However, tippers are not used in this study, and are hence not discussed further.

Apparent resistivities and phases (or impedances) from different MT stations are combined and inverted to form images of the subsurface resistivity distribution that, due to the low-frequency, large-wavelength nature of the MT source fields and induced fields, can reach hundreds of kilometers in depth (depending on the acquisition parameters). Low-frequency signals (10 Hz to 1000 s) are generated by fluctuations in the shape of Earth's magnetosphere (which occur due to temporal variations in solar wind pressure), while relatively high-frequency signals (1 kHz to 10 Hz) are generated when worldwide lightning strikes excite the Earth-ionosphere waveguide's Schumann Resonances (Chave and Jones, 2012, Chap. 3). This reliance on naturally-occurring signals circumvents a significant source of survey design complexity and deployment overhead incurred by active source techniques like seismic imaging and ERT, and as a consequence, MT data can be collected rapidly over large swaths of land. The clear downside to this dependence on natural source mechanisms, is that the quality of the recorded signals can become dominated by source

field behavior and cultural noise, factors that are largely out of a practitioner's control.

In our survey, 12 MT stations were collected over an approximately 10 km long transect, with each station left in the field for 1/2 to 4 day long recording intervals. The high-low divide's presumed location per Fackrell (2016) (see Fig. 1-1) passes through the neighborhood of Holualoa, limiting the survey's southwestern extent. Each station consisted of a Phoenix brand MTU-5A MT data logger deployed alongside Phoenix MTC-150 coil magnetometers, and Phoenix PE5 non-polarizing electrodes. Only horizontal magnetic fields were recorded (precluding the use of tippers), because the shallow topsoil layer encountered throughout the survey area was truncated by hard rock.

2-2 Data processing

The term data processing refers to the act of estimating the components of the impedance tensor (along with their uncertainties), from time domain measurements of **E** and **H**. For a single frequency, short-time Fourier transforms of the time-domain electric and magnetic field recordings constitute observations of the outputs and inputs, respectively, in a linear system defined by equation (2-1). Although ordinary least squares can then be used to estimate **Z** from equation (2-1), this approach often yields biased estimates (Chave and Jones, 2012, Chap. 5.1). So-called robust processing algorithms employ estimators rooted in the field of robust regression, which aim to mitigate the effects of 'unusual' data on the final impedance and uncertainty estimates. The 'quality' of the impedance estimates can be further improved using remote-reference processing schemes, e.g. (Gamble et al., 1979), in which each station is processed with data acquired simultaneously from a remote station, with the aim of reducing bias in the final impedance estimates.

The raw data from each station collected in our study, are processed using software implementing one of two different robust processing algorithms. Nine of the collected stations (PAL[0–2], PAL4, PALB0, and PALB[2–5]; see Fig. 2-1 for station names) are processed using a least trimmed squares algorithm described in (Jones and Jödicke, 1984), implemented using the Phoenix geophysics brand SSMT-2000 data processing toolchain (bundled with the field instruments). The remaining three stations (PALB3, PAL5, and PAL6) exhibit particularly noisy impedance estimates after processing with SSMT-2000, and are subsequently re-processed using the bounded influence algorithm described in (Chave and Thomson, 2003) and (Chave and Thomson, 2004), implemented using their Bounded In-

fluence Remote Reference Processing program (Chave, 2004). Remote reference processing schemes are applied in both cases, using data from a dedicated station located in the Big Island's Humu'ula saddle region (see Fig. 1-1 inset) as the remote reference data. The final products after applying these data processing workflows, are estimates of the impedances for each station.

2-3 Inversion strategies

With estimates of the station impedances in hand, from hereon referred to as the data, the task shifts to inferring the resistivity model that best describes the data (a type of geophysical inverse problem). One approach to solving the inverse problem is to cast it as a deterministic, non-linear constrained optimization problem, under the assumption that the data are independent and normally distributed. The Occam's Inversion (Constable et al., 1987), is one such implementation of this strategy, in which the objective function is a scalar valued function R, representing the net spatial roughness (in a first spatial derivative sense) of the model vector $\mathbf{m}$, whose elements correspond to the base-10 logarithms of the resistivities in a subsurface discretization mesh. Under the constraint that $\mathbf{m}$ must yield synthetic data that fit the vector $\mathbf{d}$ of 'observed' data, within an acceptable upper misfit bound χ^2_*, this formulation is designed to yield estimates of $\mathbf{m}$ that are maximally 'smooth' for a given χ^2_*, a property that while perhaps lacking realism (given that in highly stratified environments like sedimentary basins, the smoothness supposition is clearly violated), helps to constrain what is in general an ill-posed problem. The equivalent unconstrained minimization problem representing Occam's inversion, is given by

$$\min_{\mathbf{m}} \quad \|\mathbf{Rm}\|^2 + \mu^{-1}\left[\|\mathbf{W}(\mathbf{d}-\mathbf{F}(\mathbf{m}))\|^2 - \chi^2_*\right] \tag{2-3}$$

where the vector-valued forward function $\mathbf{F}$, represents a physics-based simulation that returns synthetic MT data, $\mathbf{W}$ is a diagonal matrix of reciprocals of the standard deviations of the data, μ is the Lagrange multiplier, and controls the degree of smoothness in the final estimate, and $\mathbf{R}$ is a bi-diagonal first difference matrix, such that the first term in the objective function constitutes the net spatial roughness of $\mathbf{m}$, and the solutions to (2-3) simultaneously minimize both misfit and roughness.

So-called Bayesian inversions offer an alternative approach to the optimization problem formulation, that seeks to estimate the posterior density function $f_{\mathbf{M}|\mathbf{D}}$, of the model

vector $\mathbf{M}$, given observations from a data distribution $\mathbf{D}$ (where the uppercase notation is used to emphasize that the quantities are random variables). Using Bayes' rule (Bertsekas and Tsitsiklis, 2008, Chap. 3.6), the posterior density function is given by

$$f_{\mathbf{M}|\mathbf{D}}(\mathbf{m}|\mathbf{d}) = \frac{f_{\mathbf{D}|\mathbf{M}}(\mathbf{d}|\mathbf{m}) f_{\mathbf{M}}(\mathbf{m})}{\int_{\Omega} f_{\mathbf{D}|\mathbf{M}}(\mathbf{d}|\mathbf{x}) f_{\mathbf{M}}(\mathbf{x}) d\mathbf{x}}, \tag{2-4}$$

where $f_{\mathbf{D}|\mathbf{M}}$ is the likelihood function, $f_{\mathbf{M}}$ is the prior distribution (also known simply as the prior), and Ω is the set of all possible values for $\mathbf{M}$. The prior is a key feature of the Bayesian approach, representing the practitioner's a priori belief in the likely parameter values, and can be used to incorporate regularization and domain specific knowledge into the model. The posterior is thus constructed by assuming a conditional distribution for the data (given parameter values), and a marginal distribution for the parameters, providing us with a method of forming parameter estimates that explicitly encode our assumptions about the overall system.

Point estimates of $\mathbf{M}$ can be obtained from equation (2-4), for example by estimating the posterior's median or maxima. This probabilistic formulation of the inverse problem also permits quantification of the uncertainty in the parameter estimates, through estimation of the posterior density function. However, interrogating the posterior by directly evaluating equation (2-4) is only computationally tractable for low-dimensional parameter spaces, due to the numerical integration required to evaluate the denominator. Samples of the posterior, drawn using a Markov chain Monte-Carlo (MCMC) sampling algorithm, can alternatively be used as a proxy for equation (2-4) (Gelman et al., 1995). In this approach, the computational effort is spent largely on repeated evaluation of $\mathbf{F}$, encountered during evaluation of the likelihood function. Because accurate representation of the posterior density generally requires a large number of draws from the posterior distribution (thousands or more, depending on the sampling algorithm used), MCMC sampling for complicated physical models remains computationally challenging in two or three spatial dimensions.

In this study, we present the results from both deterministic and probabilistic inversions of our MT data. The 2D, deterministic inversions are carried out using MARE2DEM (Key and Ovall, 2011) (Key, 2016), a dual grid, finite element, MT modeling program, that implements Occam's inversion. The data-set is restricted to the band 1.02 Hz to 780 Hz, as data quality outside this range is highly variable. Data from PALB5 is especially corrupted with cultural noise (due to its close proximity to powerlines, homes, and a highway), so we restrict it to the range 79 Hz to 450 Hz. The survey line used for constructing the model

is chosen from a least-squares fit to the station locations, yielding a trend of 19.5° North-East. We then rotate the impedance data from each station using subroutines from the Python library MTpy (Krieger and Peacock, 2014) (Kirkby et al., 2019), so as to make H_y parallel with the survey line. An 8% error floor is ascribed to the data, in order to prevent over-fitting, and a $100\,\Omega\,\mathrm{m}$ half space is used as the initial model. The data in the 2D inversion is a vector of base-10 logarithms of xy and yx apparent resistivities, along with xy and yx phase, and the parameter vector is the base-10 logarithm of the resistivities of the parameter mesh elements.

In order to identify unconstrained regions of our 2D resistivity image, the data-set used in the 2D inversion was also inverted in 1D at each station location, using Bayesian MT inversion software that we have developed. 1D inversions were chosen over 2D inversions for this task, because uncertainty quantification (UQ) using the 2D forward function was deemed to be far less computationally tractable. Noting the sparse distribution of wells in the high-level aquifer, these 1D inversions were additionally motivated by a desire to supplement well data used in groundwater modeling workflows, with electrical resistivity profiles generated from MT data. In these modeling workflows, head measurements from wells in the model domain are often considered as the data in an inverse problem seeking the subsurface distribution of hydraulic conductivity (termed model calibration in the groundwater modeling parlance). Because hydraulic conductivity is a crucial parameter for accurately modeling groundwater flow, the sparsity of wells in the high-level aquifer contributes directly to a diminished understanding of how high-level water flows through the subsurface. Towards overcoming this deficiency, we envisioned a data collection procedure in which a single MT station acts as a *virtual water well*, in that a head extracted from a 1D inversion result, could be used as a groundwater model calibration input. This would provide hydrologists with a method of rapidly and non-invasively collecting time-series observations of head at arbitrary locations on Earth's surface.

These inversions were carried out using software that we have developed in the Stan modeling language (Carpenter et al., 2017). In the 1D model, the Earth is represented as a series of discrete, homogeneous layers. Because the squared misfit term in equation (2-3) is proportional to the log-posterior density of a Gaussian distribution for $\mathbf{d}$, we assigned a Gaussian distribution to the data in the 1D model as well, so as to mimic the Occam formulation. Similarly, we chose a prior on $\mathbf{m}$ that assigns a Gaussian distribution to the parameter differences, so as to mimic the roughness term $\|\mathbf{Rm}\|^2$, in equation (2-3). The

overall statistical model is thus given by

$$\begin{aligned}
\mathbf{D}|(\mathbf{M}=\mathbf{m}, S=s) &\sim \mathcal{N}(\mathbf{F}(\mathbf{m}), s\mathbf{W}^{-1}) \\
m_i &= \log_{10}(\rho_i) \quad 2 \leqslant i \leqslant N_{\mathbf{m}} \\
\mathbf{W}^{-1} &= \operatorname{diag}(\sqrt{\operatorname{Var}(\mathbf{D})}) \\
M_i|(A=\alpha) &\sim \mathcal{N}(m_{i+1}|\alpha\sqrt{h_i}) \quad 2 \leqslant i \leqslant N_{\mathbf{m}} \\
S, A &\sim \Gamma(2, 0.1)\,,
\end{aligned} \tag{2-5}$$

where h_i and ρ_i denote the thickness and electrical resistivity (respectively) of the i^{th} layer in the Earth model, S acts as a naive method of incorporating uncertainty in the estimate of $\mathbf{W}$, and A accounts for uncertainty in the amount of smoothness that should be preferred (similar to the Lagrange multiplier μ in equation (2-3)). The implementation of the forward function $\mathbf{F}$, follows the total field impedance solution procedure described in (Iskander, 2013, Chap. 5.5). The $\Gamma(2, 0.1)$ distribution for S and A, is chosen for its zero avoiding property, preventing the sampler from taking draws near the singularity in the posterior density function at $S = A = 0$. Previous studies applying models similar to equation (2-5) for inverting MT data using an MCMC sampler, include (Guo et al., 2011) and (Conway et al., 2018). Noting that in a 1D Earth model, inversions performed on the xy and yx components of $\mathbf{Z}$ have identical results, we applied this inversion procedure to the xy and yx apparent resistivities and phases at each station, as a naive means of assessing the dimensionality of the subsurface.

2-4 Magnetotelluric inversion results

To first order, the 2D inversions from MARE2DEM (Fig. 2-1) exhibit a layered resistivity distribution with four distinct 'zones':

1. A shallow resistive zone extending to a few hundred meters below the surface of the Earth, exhibiting resistivities on the order of 10 kΩ m. These values are consistent with the values for dry, Hawaiian basaltic rock, as described by (Pierce and Thomas, 2009).

2. A mid-depth resistive zone, slightly more electrically conductive than the top zone and clustered around MSL, exhibiting resistivities between 1 kΩ m down to around

300 Ω m. The values at the high end of this range (from approximately 1 kΩ m to 600 Ω m) are consistent with freshwater saturated Hawaiian basaltic rock (per laboratory measurements published in (Rai and Manghnani, 1981, Table 2), and borehole water-level observations that were correlated with an MT survey by (Pierce and Thomas, 2009)).

3. A deep, electrically conductive zone beginning at around 1 km below mean sea level (below mean sea level (BMSL)), and exhibiting resistivities below 100 Ω m. These values are mostly consistent with saltwater-saturated Hawaiian basalts (167 Ω m to 41 Ω m (Rai and Manghnani, 1981, Table 2)), with some anomalies falling below this range.

4. A resistive zone truncating the conductive zone, beginning at around 2 km BMSL, with resistivities on the order of hundreds of Ω m.

The deep conductor dips slightly westward, and in the vicinity of station PALB2, rises to around MSL, where it takes on low resistivities on the order of 1 Ω m, far below the minimum resistivity of 40 Ω m reported by Rai and Manghnani (1981) for seawater saturated basalt. At the image's western limit, where the survey line intersects the inferred location of the high-low divide (between PALB5 and PAL6, see Fig. 1-1), the conductor rises abruptly to a few hundred meters below MSL. Again, low resistivity values on the order of a few Ω m can be observed at this depth-level. To the immediate east of this feature, a narrow resistor dips sharply to the east (dark blue color).

The 1D inversion results throughout the line (Fig. 2-2) generally agree with depth profiles extracted from the 2D inversion at the same locations. The xy and yx component inversions in most cases also agree with one another, suggesting that the Earth along the survey line has a layered structure at most locations. Deviations between these three profiles are strongest near the high low divide's indicated location, likely due to an increase in subsurface heterogeneity that invalidates the 1D modeling assumptions. The posterior resistivity distributions are well constrained above approximately 1.5 km BMSL, below which the distributions become significantly more diffuse.

To assess the fit of the 1D results to the data, we compare simulations from the posterior predictive distributions for apparent resistivity and phase, to the inverted data. The

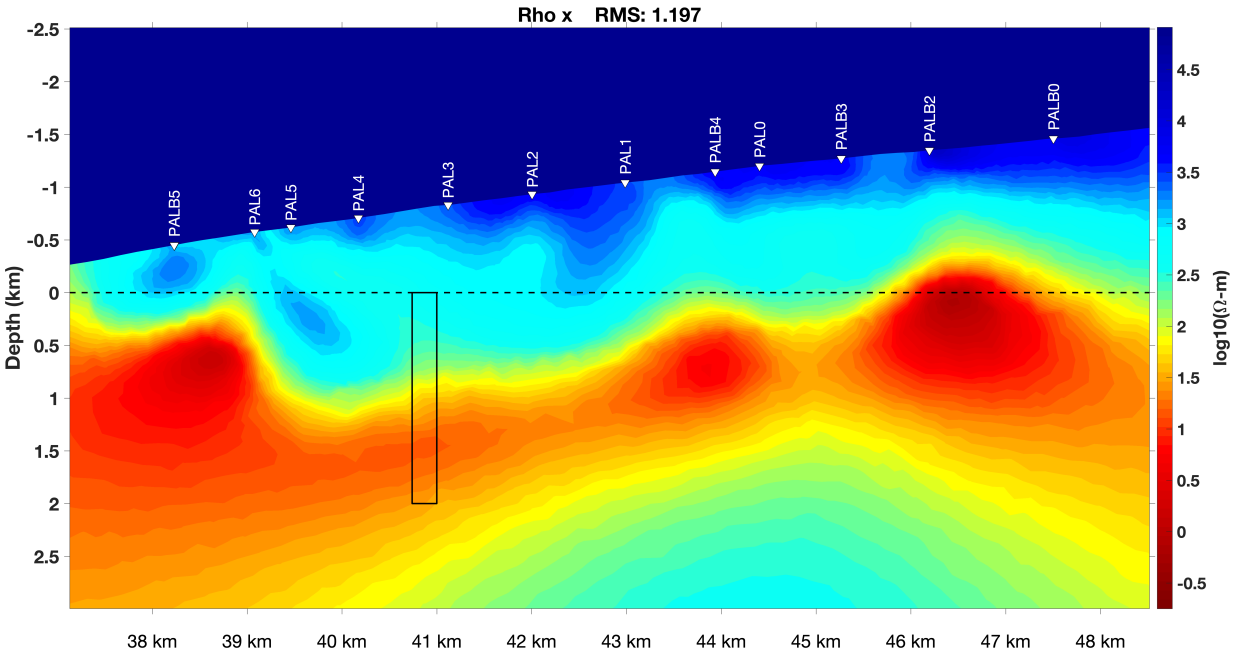


Figure 2-1: 2D deterministic inversion result from MARE2DEM. The depth scale is relative to mean sea level, and the along profile distances are in easting relative to the western extent of the mesh. The color scale has been manually adjusted so that dark-blue colors correspond roughly to the resistivity range for dry Hawaiian basalts (following Pierce and Thomas (2009)), light-blues to the range for Hawaiian basalts saturated with fresh water (following Rai and Manghnani (1981), and Pierce and Thomas (2009)), and reds to the range for seawater saturated Hawaiian basalts (following Rai and Manghnani (1981)). The divide structure's presumed location, based on heads observed in neighbouring wells (Fig. 1-1), lies between stations PALB5 and PAL6. The black rectangle at around $41\,\mathrm{km}$ along profile, indicates the region in which the salinity and porosity estimation procedure is performed.

posterior predictive distribution (Gelman et al., 1995), defined as

$$f_{\tilde{\mathbf{D}}|\mathbf{D}}(\tilde{\mathbf{d}}|\mathbf{d}) = \int_{\Omega} f_{\mathbf{D}|\mathbf{M}}(\tilde{\mathbf{d}}|\mathbf{m}) f_{\mathbf{M}|\mathbf{D}}(\mathbf{m}|\mathbf{d})\, d\mathbf{m}\,, \tag{2-6}$$

is the distribution of unobserved data $\tilde{\mathbf{D}}$, as predicted by the model's fit to the observed data $\mathbf{D}$. If the observed data appear to come from the same distribution as the predicted data, then we can conclude that the model is a reasonable representation of the data (i.e. it fits the data well). The posterior predictive distributions (see A) of apparent resistivity and phase, show good agreement with the data. The xy and yx apparent resistivities and phases generally fall within the 95% central interval (CI)s, while simulations from

the posterior predictive have similar shapes to the data distributions. This suggests that equation (2-5) is an adequate model for the data.

To test the hypothesis that the narrow, sharply dipping resistor at the western end of Fig. 2-1 is associated with a vertically-oriented, highly-resistive body (as opposed to an anomaly in the inversion result, for example due to spatial aliasing), we prepare a synthetic resistivity model containing the feature in question, and then invert the synthetic MT data generated from that model. In the synthetic model, we place a 100 m thick, vertically-oriented, 10 kΩ m resistive body at the approximate location of the dipping resistor in Fig. 2-1. This synthetic body represents a dense low-permeability structure, such as a small complex of dikes, or perhaps an unusually massive lava flow draped over a cliff. The synthetic body separates a 41 m thick low-level aquifer to its west, from a 1.04 km thick high-level aquifer to its east. The synthetic aquifers are both basal, and have a resistivity of 600 Ω m (consistent with the values for freshwater-saturated rock reported by Rai and Manghnani (1981) and Pierce and Thomas (2009)). Seawater-saturated basalt is represented in the model with a resistivity of 10 Ω m (slightly below the range presented in (Rai and Manghnani, 1981)), while dry basalt is represented with a resistivity of 3 kΩ m (an order of magnitude below the range for dry basalts reported by Pierce and Thomas (2009), so as to provide some resistivity contrast with the dike complex). The station locations, model topography, and bandwidth of inverted data, are identical to those used to prepare the field data inversion seen in Fig. 2-1. To make the synthetic data more representative of electromagnetically noisy field data, white Gaussian noise with 4% relative standard deviation is added to the simulated apparent resistivities and phases prior to inversion. This is less 'noisy' than the error floor we assigned to the field data, hence the synthetic inversion result represents an upper limit on the resolvability of the synthetic features.

The resultant resistivity image, shown in Fig. 2-3, has many features that are to first order, visually similar to those seen in Fig. 2-1. Most notable are the imprints of the synthetic high-level aquifer and dike complex. A vertically oriented resistive body straddles the western extent of the high-level aquifer, and lies to the immediate east of the dike complex. This imprint is vertically oriented, whereas the associated feature in Fig. 2-1 dips east. To the east of this feature (directly beneath PAL4), the dry zone appears to subside and 'interfinger' with the high level aquifer, with a shape resembling the resistive zone between PAL2 and PAL1 in Fig. 2-1. This feature, which has no corresponding synthetic structure, is an anomaly that we believe may be due to spatial aliasing of the dike complex (which

compared to the length scale of the MT source mechanisms, is a small wavelength feature that may require a denser station spacing to properly resolve). Between these two features, the base of the high-level aquifer has an upward 'kink', resembling the upward kink in the mid-range resistive zone of Fig. 2-1. The sharp interface between fresh- and saltwater-saturated rock in Fig. 2-3 is readily identified, although the smooth nature of the inversion algorithm makes its absolute location discernible to within only a few hundred meters. Similarly, although the high-level aquifer's head is clearly present in Fig. 2-3, accurately extracting its location is not possible.

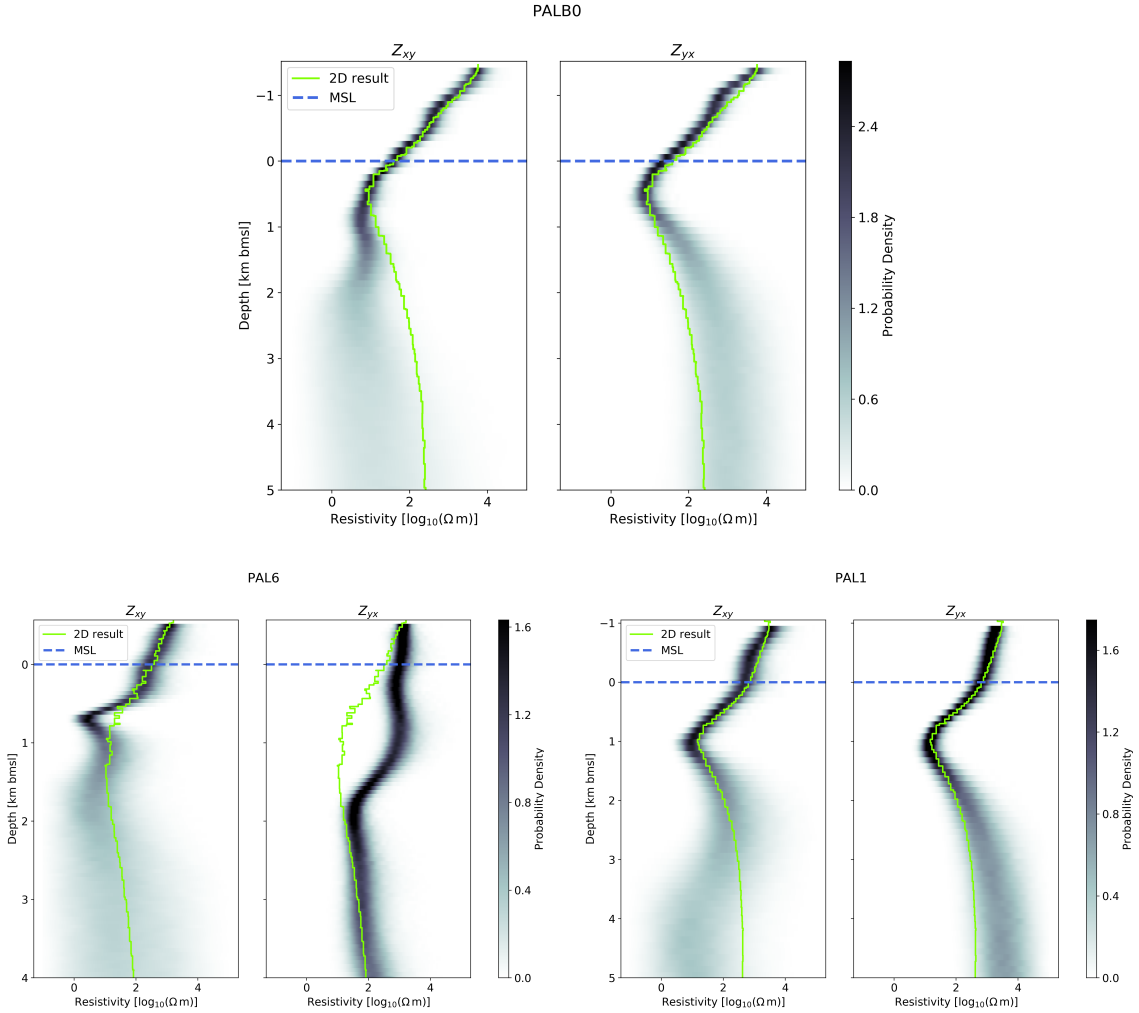


Figure 2-2: Gaussian kernel density plots of the xy- (left panels) and yx- (right panels) component posterior resistivity distributions at stations PALB0, PAL6, and PAL1. For a single depth, the darkness of each cell corresponds to the degree of belief in the parameter taking on that value, with darker colors representing greater confidence. Profiles taken from the model shown in Fig. 2-1 are overlain in green, and in general agree reasonably well with the posteriors. At PAL6, the posteriors diverge from one another, and also from the green profile, likely due to an increase in subsurface heterogeneity that violates the 1D modeling assumptions. The distributions become significantly more diffuse at approximately $1\,\mathrm{km}$ BMSL, suggesting that the resistivity models are poorly constrained below that point.

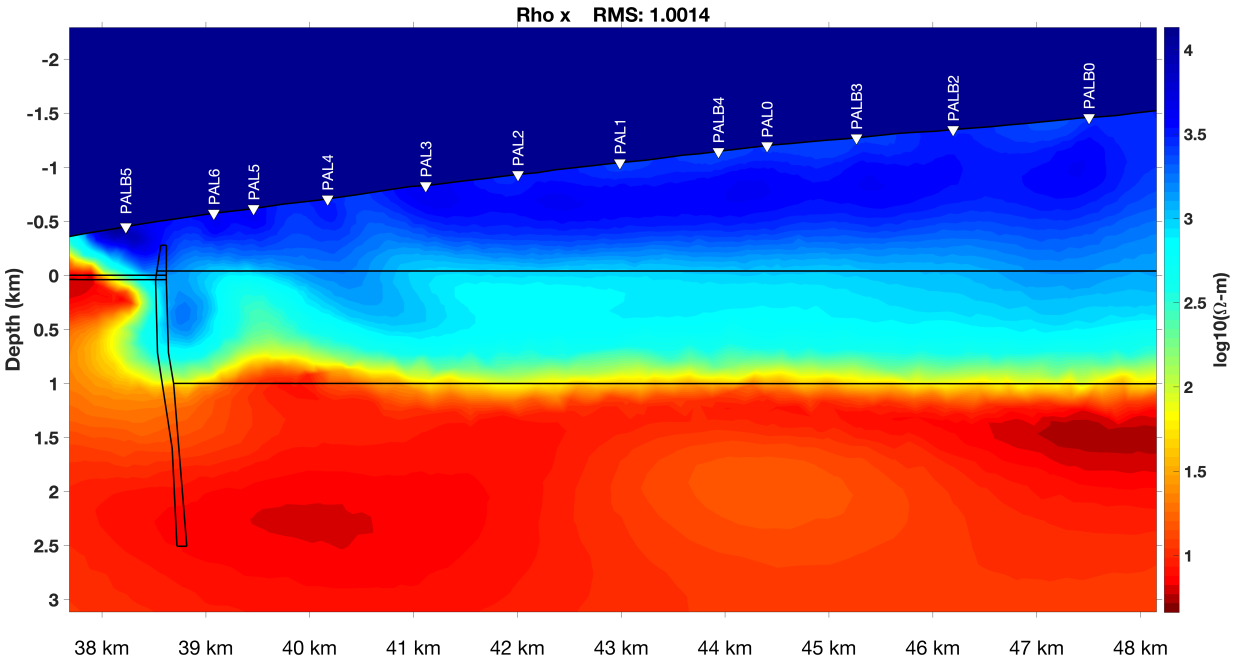


Figure 2-3: Model generated by inverting synthetic MT data simulated from a synthetic resistivity model. The interfaces in the synthetic model are overlain in black. The synthetic model features an approximately $100\,\mathrm{m}$ thick, vertically oriented, high resistivity 'sheet', separating a thin coastal aquifer from a thick basal aquifer. The vertical sheet's resistivity was set to $10\,\mathrm{k\Omega\,m}$ to represent a nearly impermeable intrusive feature, the dry zone was set to $3\,\mathrm{k\Omega\,m}$ (representative of dry basalts), the freshwater zones were set to $600\,\Omega\,\mathrm{m}$ (at the low end of the freshwater saturated basalt resistivities reported by Rai and Manghnani (1981, Table 2)), and the saltwater zone was set to $10\,\Omega\,\mathrm{m}$ (slightly more conductive than the seawater saturated basalt resistivities report by Rai and Manghnani (1981, Table 2)). To facilitate comparison, the simulated frequencies were chosen to be identical to those used in the field data inversion (Fig. 2-1).

Chapter 3

Hydrogeological parameter estimation

3-1 Salinity and porosity

Due to the non-unique nature of the MT inverse problem, the importance of quantifying uncertainty during the interpretation stage cannot be understated. In our analyses, we should consider not only the uncertainty in our resistivity estimates, but also in the hydrological inferences we draw from them. Recall that one of our goals in this study, is to identify fresh and saline groundwater in the high level aquifer, along with the depths at which they occur. One method of constructing quantitative estimates of both salinity and porosity (which together can be related to bulk resistivity), along with their uncertainties, is to again use equation (2-4) to relate the estimated resistivities to the hydrological parameters of interest. Consider a fully-saturated formation (e.g. an aquifer). Let Φ_i denote the porosity of the volume of rock represented by the i^{th} cell of a discretization mesh, and R_i^w the electrical resistivity of the saturant in that cell, where the uppercase notation again denotes a random variable. We will use Archie's law (Archie, 1942), originally developed in sedimentary environments, but demonstrated by Rai and Manghnani (1981) to be applicable to Hawaiian basalts, as our physical model relating bulk resistivity, porosity, and fluid resistivity. Denoting an estimate of the bulk resistivities in the mesh as $\mathbf{M}$ (in our case, the 2D inversion result from MARE2DEM), we form a distribution for $\mathbf{M}$ following Chave and Jones (2012, Chapter 8), who noted that for a fixed value of the Lagrange multiplier μ, equation (2-3) implies a Gaussian distribution for $\mathbf{M}$. Denoting the Jacobian of the MT

forward functional as $\mathbf{J}$, the model covariance $\mathbf{C_M}$ is given by

$$\mathbf{C_M} = \left(\mu \mathbf{R}^T \mathbf{R} + \mathbf{J}^T \mathbf{W}^{-1} \mathbf{J}\right)^{-1}. \tag{3-1}$$

Next, we formulate priors for the parameters Φ_i and R_i^w. Because porosities are restricted to the set $(0, 1]$, we determine the beta distribution to be a natural choice of prior for Φ_i, and introduce a dataset of observed porosities $\hat{\Phi}_k$ (which are independent of the mesh), in order to constrain the Beta shape parameters A and B. Φ_i and $\hat{\Phi}_k$ are both assigned the same distribution. Denoting salinity (in parts per thousand (PPT) of total dissolved solids) as $\hat{C}_i$, we form the fluid resistivity prior by first assuming a linear relationship between R_i^w and $\hat{C}_i$, which allows us to instead place a prior on salinity. Noting that salinities are strictly positive, and assuming that they should generally fall between tap-water and seawater, we choose a Log-normal distribution for $\hat{C}_i$, with the location and shape parameters μ_c and σ_c, determined by setting quantiles for the salinities of tap-water and seawater, and then using an algebraic solver to determine corresponding values of μ_c and σ_c. The overall statistical model is given by

$$\begin{aligned}
\mathbf{M}|(\mathbf{\Phi} = \phi, \mathbf{R}^w = \rho_w) &\sim \mathcal{N}(\mu_{\mathbf{M}}(\phi, \rho_w), \mathbf{C_M}) \\
\mu_{\mathbf{M}} &= -n \log_{10}(\Phi) + \log_{10}(\rho_w) \\
\Phi_i, \hat{\Phi}_k|(A = a, B = b) &\sim \text{Beta}(a, b) \\
\log_{10}(R_i^w)|(\hat{C}_i = \hat{c}_i) &= -a_w \log_{10}(\hat{c}_i) + b_w \\
\hat{C}_i &\sim \text{Lognormal}(\mu_c, \sigma_c) \\
A, B &\sim \text{U}(0, \infty)\,,
\end{aligned} \tag{3-2}$$

where $\mu_{\mathbf{m}}$ is the base-10 logarithm of Archie's law (Archie, 1942) with cementation exponent n. Assuming that the lithology throughout the survey area is constant, n can be set to a constant value.

We apply this porosity/salinity estimation method to a rectangular segment of the MARE2DEM inversion images (see Fig. 2-1). The size of the segment is made small, so as to minimize run-time and facilitate rapid experimentation with the statistical model. Because the form of Archie's law used in equation (3-2) assumes that the formation is fully saturated, the segment only includes cells that are beneath sea level. To construct the dataset $\hat{\Phi}$, porosities from drill cores taken in Kilauea and the Humu'ula Saddle are aggregated from Revil et al. (2016, Table 1), and Revil et al. (2018, Tables 1 & 2). Assuming that

the study site lithology is entirely basaltic, the value $n = 2.57$ is also adopted from Revil et al. (2018). The slope and offset parameters a_w and b_w, which are used to define a linear relationship between salinity and fluid resistivity, are determined from a least-squares fit to synthetic samples of resistivity and salinity, generated by passing randomly drawn resistivity samples through subroutines from the Gibbs Sea Water Toolkit (McDougall and Barker, 2011), assuming a temperature of 25 degrees Celsius, and atmospheric pressure.

Samples from the posterior are drawn using software we have written in the Stan modeling language (Carpenter et al., 2017). To avoid any inaccuracies introduced by the imposed linear relationship between salinity and fluid resistivity, the final salinity estimates are generated by converting the sampled fluid resistivities to absolute salinities using The Gibbs Sea Water Toolkit, again at a temperature of 25 degrees Celsius, and atmospheric pressure. The constants μ_c and σ_c are determined by setting the 0.01 and 0.99 quantiles of $\hat{C}_i$ to 0.07 and 35, respectively. In units of PPT of total dissolved solids, these values correspond to the tap-water and seawater solutions used in (Rai and Manghnani, 1981, Table 1). Finally, profiles of salinity and porosity, taken midway between the western and eastern limits of the inverted segment (the black rectangle in Fig. 2-1), are prepared from the draws (shown in Fig. 3-1).

3-2 Results

The posterior salinity profile (Fig. 3-1) increases with depth, taking on its largest value at around 1.25 km BMSL. The posterior porosity profile also shows an increase at this depth, but otherwise shows little variation. At mean sea level, the median salinity is around 1.0 PPT, and the mean salinity is around 2.0 PPT. With increasing depth, these central tendencies for salinity approach, but do not reach seawater salinity, despite the 2D and 1D resistivity estimates reaching the seawater range in (Rai and Manghnani, 1981, Table 2). Mean salinity throughout the profile is always greater than the median, and the 95% CI for salinity is very wide throughout the profile. Draws from the posterior predictive distribution (Fig. 3-2) show some significant deviations from the MARE2DEM estimate, but generally follow its overall shape. The most noticeable discrepancy is the lack of a distinct peak near saltwater values, although some flattening of the curves is seen in this range. This suggests that the statistical model underestimates the salinity of seawater. The draws appear to share a maximum with the MARE2DEM estimate, but are also observed to have heavier tails, especially at the high-resistivity end. The residuals

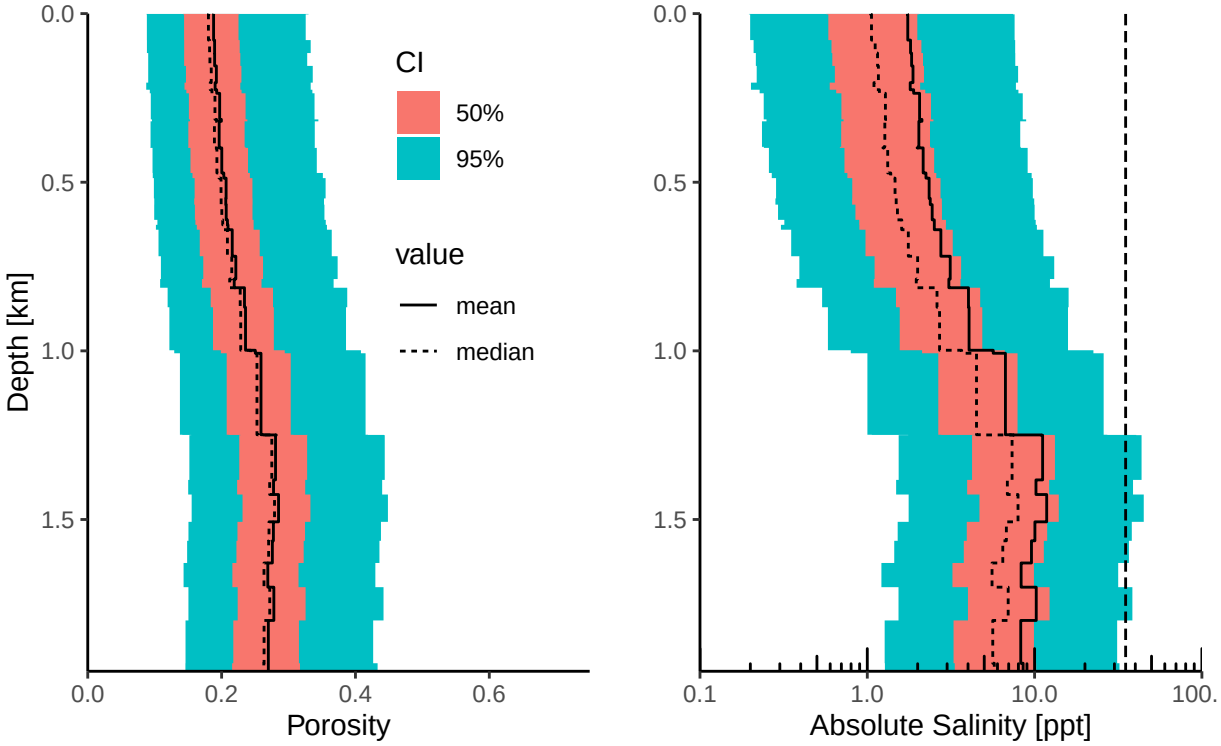


Figure 3-1: Posterior distributions of porosity (left) and absolute salinity (right), taken midway between the eastern and western limits of the rectangular segment in Fig. 2-1. Depths are in meters below mean sea level. The vertical dashed line in the salinity plot, indicates the salinity of seawater. Both the salinity and porosity distributions take on their maximum values between 1 km and 1.5 km, with the mean of salinity approaching (but never reaching) seawater salinity.

between the draws and the MARE2DEM estimate appear to be centered about zero, and have flat spots at positive and negative $10\,\Omega\,\mathrm{m}$, likely due to the discrepancies noted at high and low resistivities.

3-3 Hydrogeological interpretation

We have prepared a resistivity image with the following interpretation overlain, shown in Fig. 3-3. The electrical resistivities observed in Fig. 2-1 are consistent with (in order of increasing depth): unsaturated rock, freshwater saturated rock, saltwater saturated rock, and freshwater saturated rock. The diffuse behavior observed in the 1D inversions at depth coincides with the deep resistor's location, suggesting that its properties are poorly constrained by the data. Although the salinity of pure seawater is only briefly

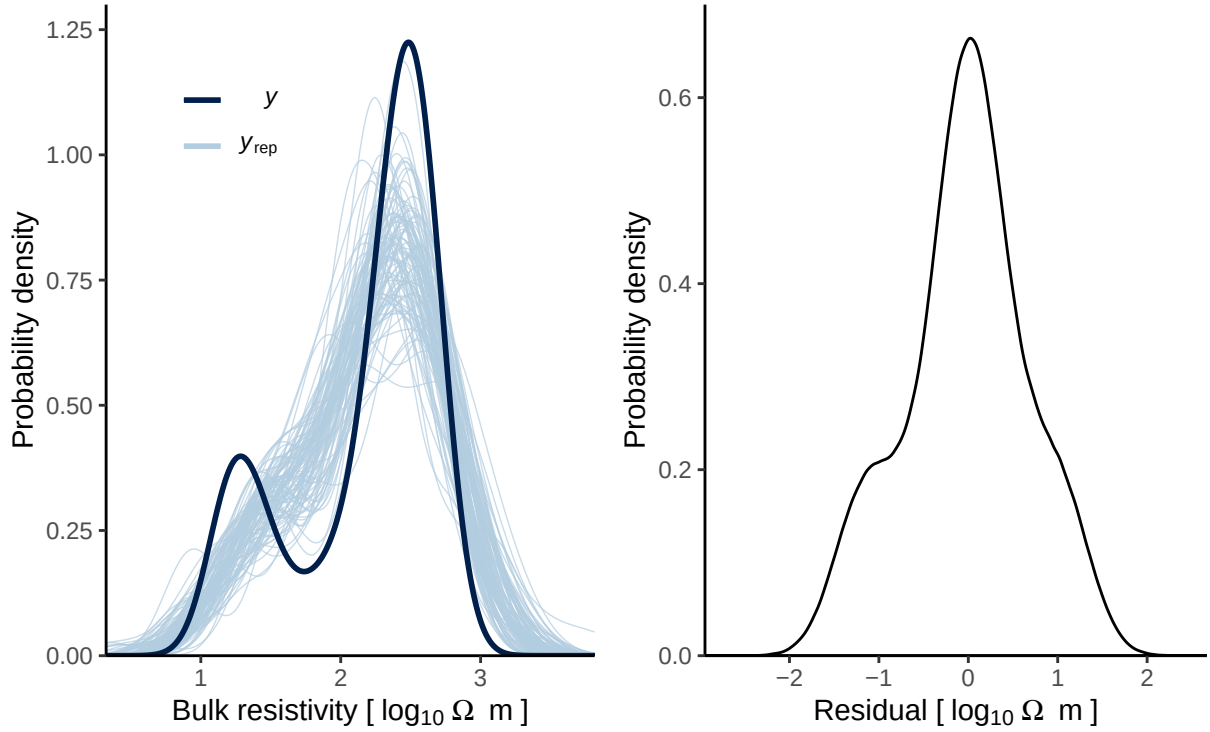


Figure 3-2: Left: Distribution of the resistivities in the inverted segment in Fig.2-1 (solid black), compared to draws from the posterior predictive distribution (transparent blue). Right: Distribution of the residuals $y - y_{\text{rep}}$. The posterior predictive draws lack the distinct low resistivity mode seen in the data, and have heavier tails.

surpassed by the salinity profile (Fig. 3-1), samples surpassing 17 PPT (the salinity at a transition zone's midpoint (Lau and Mink, 2006)) occur at roughly 1 km BMSL. This suggests that a transition from fresh to salt water occurs in the high level aquifer, at a depth of roughly 1 km BMSL (Fig. 3-3). It is not surprising that the inferred depth to saltwater coincides with the depth of investigation, since increases in conductivity yield corresponding decreases in electromagnetic skin depth (Chave and Jones, 2012). Because skin depth also decreases with increasing frequency (Chave and Jones, 2012), the depth of investigation in our results, which for an MT survey can be considered very shallow, is likely to be further limited by the narrow, high-frequency bandwidth of the data used in the inversions.

Between PAL1 and PAL5, a transition from unsaturated to freshwater-saturated rock is not immediately discernible from the images. The contact between the shallow and mid depth resistors is erratic, and not representative of a confining layer, nor a contiguous water table.

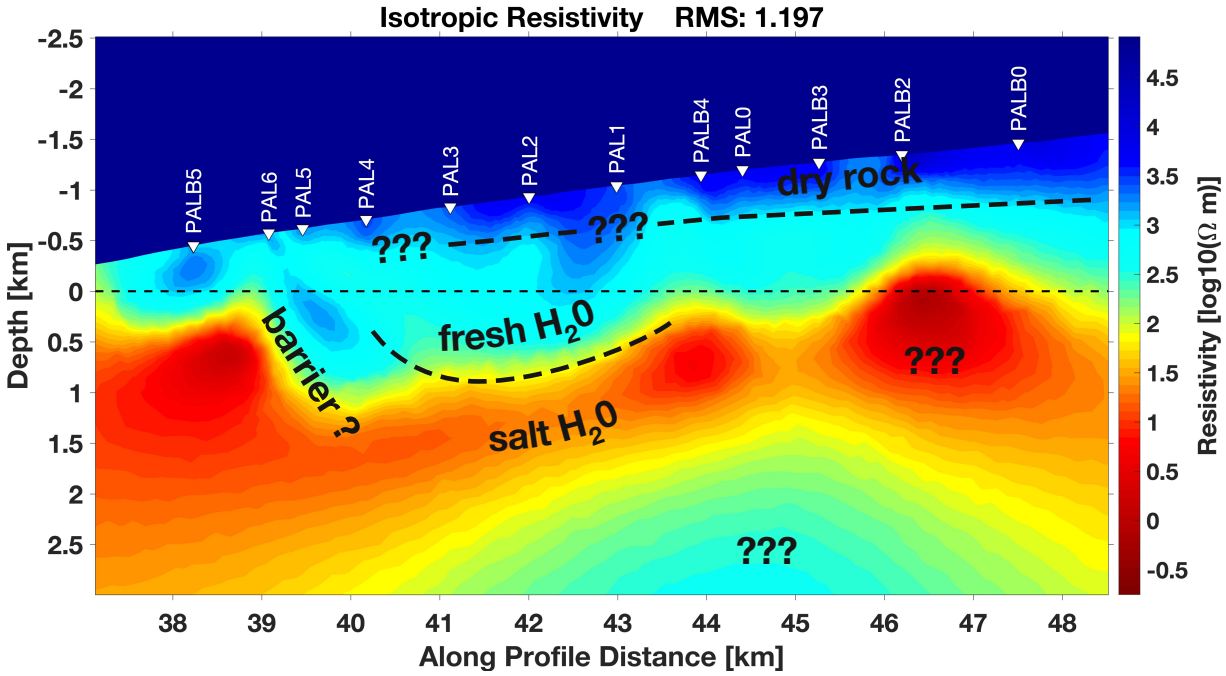


Figure 3-3: 2D inversion result, with the hydrogeological interpretations and inferred boundaries overlain.

If high-level freshwater is indeed hydraulically continuous with saltwater, then a freshwater head level, which in an unconfined aquifer coincides with the water table, can be estimated from the inferred transition zone's location. Assuming that fluid motion is predominantly horizontal, and that the high-level freshwater is truncated by seawater, substituting our inferred depth to saltwater of approximately 1 km into the Ghyben-Herzberg ratio, gives a head of 25 m. Although low for a high-level well, this head is on the order of the high-level heads observed to the immediate north of the survey line (between 40 m and 100 m AMSL (Bauer, 2003)), suggesting that our estimated depth to seawater is reasonably consistent with field observations of head.

To the east of PAL1, the transition from high- to mid-range resistivities is markedly sharper. Mid-range resistivities consistent with freshwater-saturated basalts occur primarily above sea level, which does not corroborate the interpretation of a vertically extensive basal aquifer. A 'flattening' is observed in the 1D inversions at PALB0 (Fig. 2-2), which may be representative of a shallow perched aquifer. This flattening is absent in the corresponding profiles from the 2D inversion, possibly because the Occam algorithm that was used to generate the 2D inversion, has smoothed over this feature. The shape of this contact between high- and mid-range resistivities bears some similarity to the shallow perched

aquifer in Fackrell (2016)'s Keauhou North conceptual model (Fackrell, 2016, Fig. 3.7c), in that both follow the topography.

The abrupt 'jump' in the deep conductor seen between PALB5 and PAL6, is likely a signature of the high-low divide. Beneath PALB5, resistivities consistent with freshwater saturated basalt occur at depths of up to roughly 300 m BMSL, which using the Ghyben-Herzberg ratio, yields a maximum head of 7.5 m. This inferred head is consistent with the low level heads reported by Bauer (2003). Bauer (2003) also notes an average water level of 1.5 m for a well to the immediate east of PALB5 (well name Pāhoehoe in (Bauer, 2003, Table 3)), which is on the order of (and hence agrees with) our inferred maximum head.

3-4 Discussion

It is important to note that our simplistic analyses that use the Ghyben-Herzberg ratio to infer heads from inferred seawater contacts, do not account for the presence of confining layers, for which there is evidence in both the high- and low-level systems (Bauer, 2003). Thus, while these results suggest that the inferred depths to seawater are reasonably consistent with field observations of head, they cannot serve as accurate head levels.

Contrary to the qualitative analyses, the central tendency of the estimated salinity profile does not corroborate the interpretation that the high-level freshwater body is truncated by seawater, since inspection of Fig. 3-1 shows that only a small fraction of the draws exceed seawater salinity. This may be due to deficiencies in the petrophysical model used in eqn. (3-2), which neglects the rock's surface conductivity, a quantity which Revil et al. (2016) and Revil et al. (2018) note is crucial for correct application of Archie's law in volcanic environments. Similarly, Rai and Manghnani (1981) observed variations in the Archie's law parameters as a function of salinity, which they noted could be due to surface conduction effects. Surface conduction effects are neglected in our analyses, in order to reduce the number of unknown parameters in what is already a strongly underdetermined system. The log-normal salinity prior is also somewhat ad hoc, in that it does not incorporate field observations, aside from the choice of tail weights (the values of which are somewhat arbitrary, and simply reflect soft bounds on salinity). While the positive skew of the log-normal distribution may adequately represent salinity in the freshwater zone, at higher salinities this shape may place too much mass above seawater values, forcing the posterior mode to remain at lower salinities.

Although Revil et al. (2018), from whom we adopt our value of the Archie's law cementation exponent n in equation (3-2), claim that this value of n "can be used for all volcanic rocks in Hawai'i"(Revil et al., 2018, Page 48), our study does not include actual porosity observations (due to a lack of samples) to corroborate this claim and justify our adoption of their value (recall that the observed porosities $\hat{\Phi}$ in equation (3-2) are from cores taken outside of our study site). We also suspect that the types of data used in the porosity and salinity estimation (bulk electrical resistivities from an inversion model, and independent core log porosities), are likely not informative enough to constrain the values of all three Archie's law parameters (fluid resistivity, porosity, and cementation exponent). The sensitivity of the solution to the value of of the Archie's law cementation exponent n, could however be investigated by considering a single value $\hat{n}$, taken from the literature, to be an observed data point, and then ascribing a distribution to this random variable (with parameters ideally also taken from the literature). Comparing the distributions in Fig. 3-1, which consider n to be known perfectly, to those produced when n is considered a random variable, should give some idea of the sensitivity of the posterior distribution to changes in n.

The non-identifiability of the porosity and salinity parameters in Archie's law, may also act to lower the salinity estimates at depth. Given a fixed cementation exponent, Archie's law can satisfy any value of bulk resistivity using an infinite number of porosity and fluid resistivity combinations, with decreases in either of the two resulting in a simultaneous increase in the other. The porosity and salinity estimates we have presented should thus be considered to represent bounds on the allowable values of the two parameters. Noting that in Fig. 3-1, both porosity and salinity take on their maximum values at roughly the same depths, it may be true that fixing porosity to a single value at the lower end of its distribution, will yield a simultaneous increase in the salinity profile's maximum value, and hence a greater number of samples meeting or exceeding seawater salinity. We additionally note that although very few samples meet or exceed seawater salinity, the 95% CI first exceeds a salinity of 17 PPT (representative of the transition zone midpoint (Lau and Mink, 2006)) at around 1 km BMSL, which is indeed consistent with our qualitatively inferred depth to seawater.

The seawater resistivities occurring at MSL between PAL1 and PALB0, are at present not readily explained, nor is the apparent large body of freshwater above MSL coinciding with it. It is possible that this 'uplift' in the deep conductor is indeed associated with saltwater saturated basalts, an interpretation that disagrees with the laterally contiguous

and extremely deep high-level freshwater bodies envisioned by Oki (1999) and Fackrell (2016). Because the survey line is situated on the flank of an active volcano, hydrothermal or chemical alteration to the rock in this zone leading to a decrease in its resistivity, may explain this anomaly, and would not preclude the presence of freshwater at these depths. The transport of thermal fluids from the volcano's core and/or remnant geothermal heat may also explain the resistivities in this zone, as increases in temperature are known to yield corresponding decreases in the resistivities of basalts (Pierce and Thomas, 2009).

The resistivity image prepared from our synthetic modeling test (shown in Fig. 2-3) displays features that resemble those seen in Fig. 2-1. This result agrees with the hypothesis that the divide structure is comprised of low permeability intrusives. However, the resistive 'barrier' in Fig. 3-3 dips east, which does not agree with any of the structures proposed by Oki (1999) and Fackrell (2016) (all of which either dip to the west, or are vertically oriented). We consider the results of the synthetic modeling test to be inconclusive, because they do not explore the sensitivity of the results to changes in the shapes and orientations of the synthetic structures, nor do they explore alternative conceptual models (the features of which may have similar imprints to those in Fig. 2-3, which would make these models indistinguishable from one another). Extensions to this test should consider the fracture and massive lava flow models of Oki (1999), and the Keauhou North model of Fackrell (2016, Fig. 3.7c). Experimentation with the perching formations of the latter may provide insight into the shape of the shallow resistor in Fig. 3-3.

3-5 Conclusions

Our 2D deterministic MT electrical resistivity inversion results for an approximately 10 km long transect in the Keauhou aquifer, on Hawai'i's Big Island, exhibits a layered resistivity structure, with three distinct (and reasonably well constrained) 'layers'. High electrical resistivities consistent with dry basalt occur at depths of up to a few hundred meters beneath ground surface. Low resistivities associated with seawater-saturated basalt occur at approximately 1 km BMSL. Mid-range resistivities consistent with freshwater-saturated basalt occur between the two. These observations are consistent with a high-level resource that is vertically extensive and truncated with seawater, most likely via a brackish mixing zone.

Based on 1D Bayesian MT inversions, a 2D deterministic MT inversion in the Keauhou

aquifer was deemed to be reasonably well constrained at depths up to approximately 1 km BMSL. Beneath this depth, the 1D profiles become markedly more diffuse, suggesting that deeper features are poorly constrained by the data. The depth of investigation is likely limited by the narrow, high-frequency bandwidth of the inverted data, along with the presence of a strong conductor (as strong conductors are known to limit the penetration depth of EM signals and affect imaging resolution). Although the 1D inversions appear to have been an effective method for assessing the reliably interpretable regions of our 2D inversion, we caution against haphazard interpretation of 1D MT inversion results (especially their structure), as this work has also demonstrated the well known phenomenon of an increase in subsurface heterogeneity grossly compromising a 1D resistivity model (see PAL6 in Fig. 2-2). A more comprehensive workflow should include subsurface dimensionality analyses in the vein of those described in (Chave and Jones, 2012, Chap. 6).

In both the 1D and 2D inversions, it is not possible to accurately identify the exact spatial transition location from unsaturated to saturated rock, precluding any determination of head from the MT data alone. This is likely due to the long-period, large-wavelength nature of the MT source fields (which are better suited to imaging broad structures with strong resistivity contrasts), compounded by the use of inversion routines that emphasize smooth models. The utility of using these results as direct inputs to a groundwater flow modeling workflow (for example, by using single station MT inversions as virtual water-wells approach), thus seems limited, although we anticipate that they will be useful as a means of qualitatively validating the structures in such models.

An abrupt west-to-east deepening of the supposed transition zone is consistent with the known location of a hydrological discontinuity (the high-low divide) near Mamalahoa highway. This suggests that the high-low divide's presumed location is accurate on the survey line to within a few kilometers. A narrow high-resistivity feature in this region of the image may be representative of a dense vertically-oriented structure, but bears an eastward dip that is not consistent with existing conceptual models for the divide structure.

The 95% CI of the posterior distribution of salinity (as a function of depth) surpasses a salinity of 17 PPT at a depth of approximately 1 km BMSL, a result which corroborates a 1 km depth to seawater saturated rock. In contrast, the median of the posterior does not surpass a salinity of 10 PPT, and only a small portion of the 95% CI meets or surpasses a salinity of 35 PPT (representative of seawater). This may be due to deficiencies in the petrophysical model, which does not account for surface conduction effects, and also the statistical model, which employs a salinity prior that may not be flexible enough (or may

have too much of a positive skew) to adequately represent seawater.

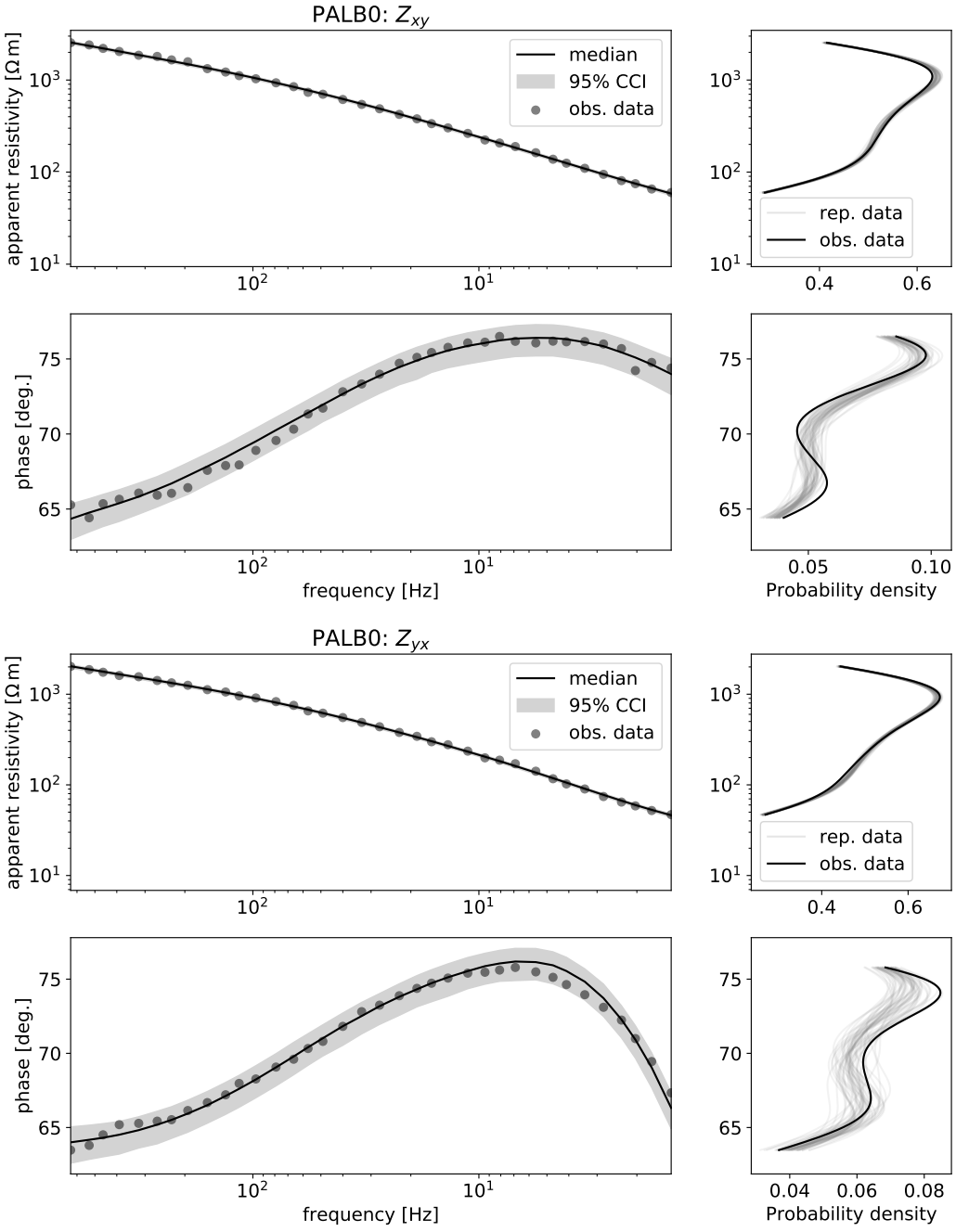
Towards further exploring the application of MT data as a direct input into a groundwater flow modeling workflow (the so-called *virtual well* methodology), it may be worthwhile to experiment with a sharp boundary 1D inversion routine, such as that described by (Conway et al., 2018), along with priors that allow resistivity gradients larger than those afforded by Occam's inversion, for example using the Horseshoe Prior described in (Piironen and Vehtari, 2016). By allowing sharper transitions in resistivity with respect to depth, these approaches may permit more accurate identification of the transition point from dry to wet rock. We note that the potential applicability of these approaches is not limited to the inversion of MT data for electrical resistivity, in that they could also be used to extend our salinity/porosity inversion methodology.

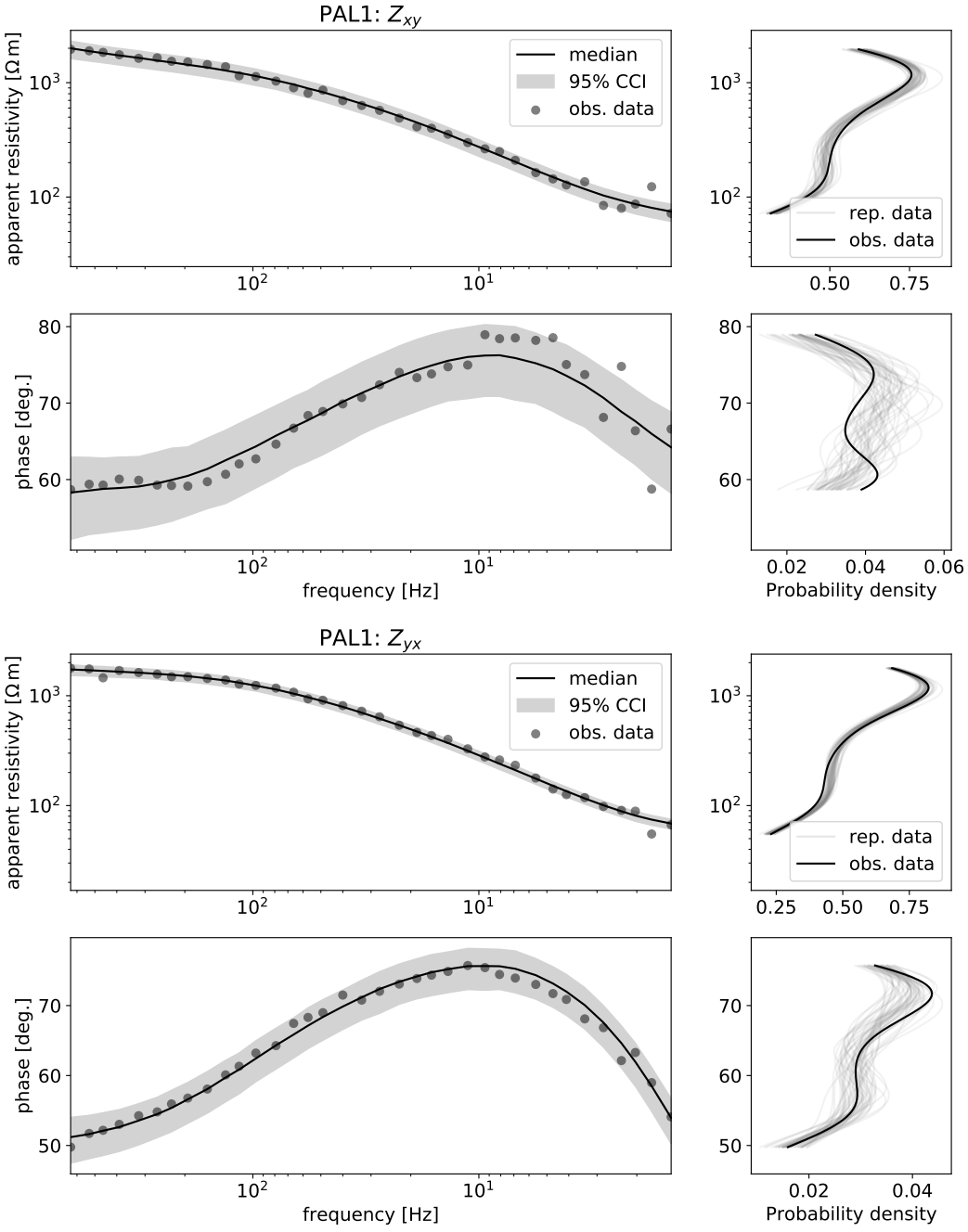
Future work on the salinity and porosity estimation methodology we have described should incorporate salinity observations from high-level wells into the prior distribution, and explore priors for salinity that have more flexible shapes. In particular, distributions with skew varying as a function of depth may do a better job of representing saltwater-saturated regions, because their masses can be constrained to always concentrate within observed upper and lower salinity bounds. Towards further exploring the hypothesis that the high-level system is indeed truncated with seawater, it may also be worthwhile to experiment with priors that explicitly encode the belief that salinity increases with depth.

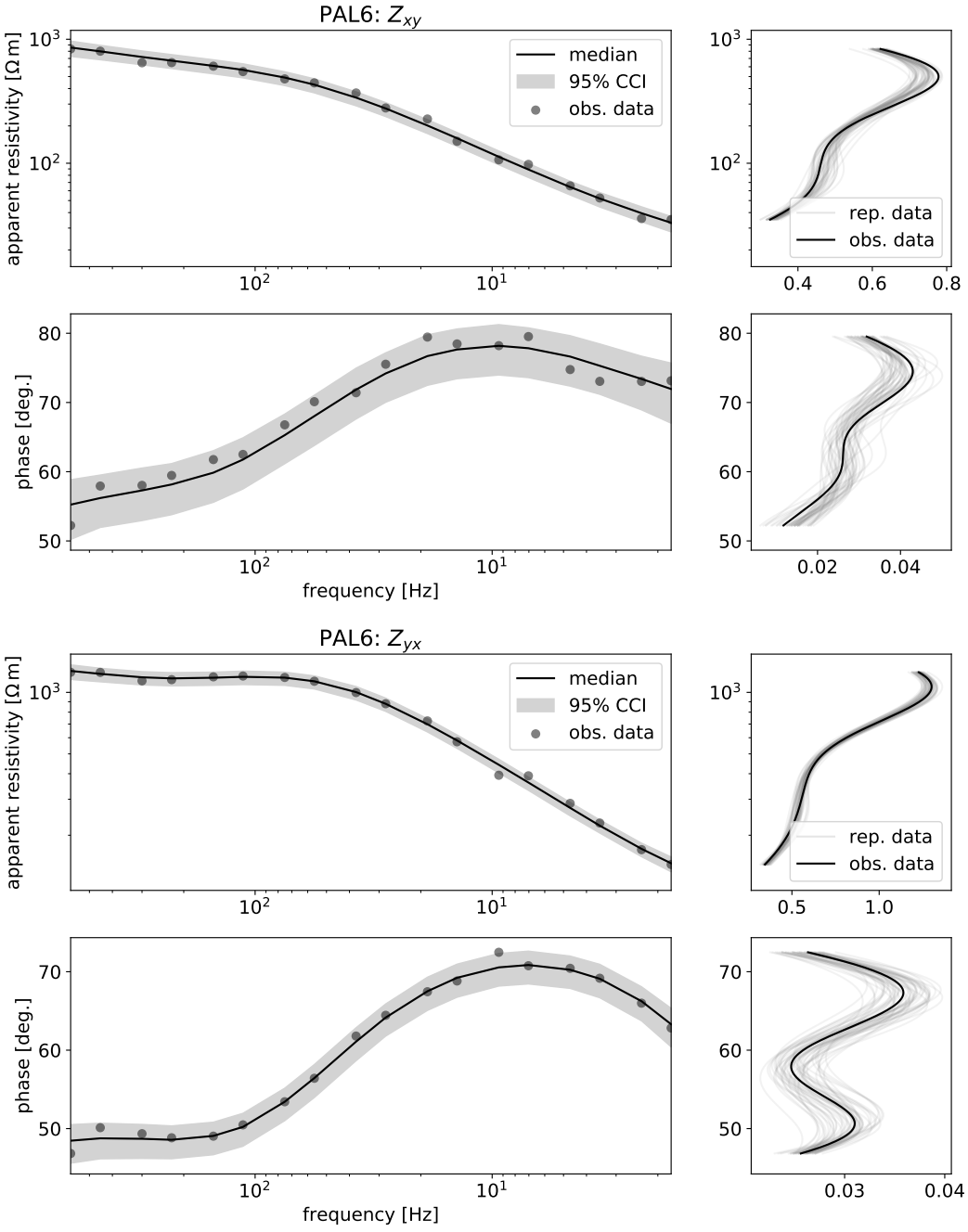
Appendix A

Posterior predictive distributions of apparent resistivity and phase

Left panels: Posterior predictive distributions of apparent resistivity and phase, corresponding to the 1D inversions shown in Fig. 2-2. Right panels: Kernel density estimates (marginalized over frequency) of the observed data (solid black), compared to draws from the posterior predictive (transparent grey). The data are mostly contained within the 95% CIs, and the posterior predictive draws appear representative of the data distributions. This suggests that equation 2-5 is an adequate model for the data.


PALB0: Z_{xy}
apparent resistivity [Ω m]
median
95% CCI
obs. data
rep. data
obs. data
phase [deg.]
frequency [Hz]
Probability density
PALB0: Z_{yx}
apparent resistivity [Ω m]
median
95% CCI
obs. data
rep. data
obs. data
phase [deg.]
frequency [Hz]
Probability density


PAL1: Z_{xy}
apparent resistivity [Ω m]
median
95% CCI
obs. data
rep. data
obs. data
phase [deg.]
frequency [Hz]
Probability density
PAL1: Z_{yx}
apparent resistivity [Ω m]
median
95% CCI
obs. data
rep. data
obs. data
phase [deg.]
frequency [Hz]
Probability density


PAL6: Z_{xy}
apparent resistivity [Ω m]
median
95% CCI
obs. data
rep. data
obs. data
phase [deg.]
frequency [Hz]
PAL6: Z_{yx}
apparent resistivity [Ω m]
median
95% CCI
obs. data
rep. data
obs. data
phase [deg.]
frequency [Hz]

Bibliography

Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME*, 146(01):54–62. SPE.

Bauer, G. R. (2003). A study of the ground-water conditions in north and south kona and south kohala districts island of hawaii, 1991-2002. Published report, State of Hawaii Department of Land and Natural Resources, Commision on Water Resources Management.

Bertsekas, D. P. and Tsitsiklis, J. N. (2008). *Introduction to Probability*. Athena Scientific, Nashua, NH, 2nd edition.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, Articles*, 76(1):1–32.

Chave, A. D. (2004). Birrp: Bounded influence remote reference processing. https://www.whoi.edu/science/AOPE/people/achave/Site/Next1.html.

Chave, A. D. and Jones, A. G., editors (2012). *The Magnetotelluric Method: Theory and Practice*. Cambridge University Press, Cambridge.

Chave, A. D. and Thomson, D. J. (2003). A bounded influence regression estimator based on the statistics of the hat matrix. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(3):307–322.

Chave, A. D. and Thomson, D. J. (2004). Bounded influence magnetotelluric response function estimation. *Geophysical Journal International*, 157(3):988–1006.

Constable, S. C., Parker, R. L., and Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *GEOPHYSICS*, 52(3):289–300.

Conway, D., Simpson, J., Didana, Y., Rugari, J., and Heinson, G. (2018). Probabilistic magnetotelluric inversion with adaptive regularisation using the no-u-turns sampler. *Pure and Applied Geophysics*, 175(8):2881–2894.

Fackrell, J. (2016). *Geochemical evolution of Hawaiian groundwater*. PhD thesis, Department of Earth Sciences, School of Ocean and Earth Sciences and Technology, University of Hawai'i at Mānoa.

Gamble, T. D., Goubau, W. M., and Clarke, J. (1979). Magnetotellurics with a remote magnetic reference. *GEOPHYSICS*, 44(1):53–68.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman and Hall Texts in Statistical Science. Chapman and Hall, London, UK, 1st edition.

Gingerich, S. B. and Oki, D. S. (2000). Ground water in hawaii. Technical report, U.S. Geological Survey. Report.

Guo, R., Dosso, S. E., Liu, J., Dettmer, J., and Tong, X. (2011). Non-linearity in Bayesian 1-D magnetotelluric inversion. *Geophysical Journal International*, 185(2):663–675.

Iskander, M. F. (2013). *Electromagnetic fields and waves*. Prentice Hall, Englewood Cliffs, N.J., 2nd edition.

Izuka, S. K., Engott, J. A., Rotzoll, K., Bassiouni, M., Johnson, A. G., Miller, L. D., and Mair, A. (2018). Volcanic aquifers of hawai'i—hydrogeology, water budgets, and conceptual models. Technical report, U.S. Geological Survey, Reston, VA. Report.

Jones, A. G. and Jödicke, H. (1984). Magnetotelluric transfer function estimation improvement by a coherence-based rejection technique. In *SEG Technical Program Expanded Abstracts 1984*, pages 51–55. The Society of Exploration Geophysicists.

Key, K. (2016). MARE2DEM: a 2-D inversion code for controlled-source electromagnetic and magnetotelluric data. *Geophysical Journal International*, 207(1):571–588.

Key, K. and Ovall, J. (2011). A parallel goal-oriented adaptive finite element method for 2.5-D electromagnetic modelling. *Geophysical Journal International*, 186(1):137–154.

Kirkby, A., Zhang, F., Peacock, J., Hassan, R., and Duan, J. (2019). The mtpy software package for magnetotelluric data analysis and visualisation. *Journal of Open Source Software*, 4(37):1358.

Krieger, L. and Peacock, J. R. (2014). Mtpy: A python toolbox for magnetotellurics. *Computers & Geosciences*, 72:167 – 175.

Lau, L. S. and Mink, J. F. (2006). *Hydrology of the Hawaiian Islands*. University of Hawaii Press.

McDougall, T. J. and Barker, P. M. (2011). *Getting started with TEOS---10 and the Gibbs Seawater (GSW) Oceanographic Toolbox*. SCOR/IAPSO WG127.

Mink, J. F. and Lau, L. S. (1993). *WRRCTR No.191 Aquifer identification and classification for the island of Hawaii: Groundwater protection strategy for Hawaii*. Water Resources Research Center, University of Hawaii at Manoa.

Oki, D. S. (1999). Geohydrology and numerical simulation of the ground-water flow system of kona, island of hawaii. Technical report, U.S. Geological Survey. Report.

Pierce, H. A. and Thomas, D. M. (2009). Magnetotelluric and audiomagnetotelluric groundwater survey along the humu'ula portion of saddle road near and around the pohakuloa training area, hawaii. Technical report, U.S. Geological Survey. Report.

Piironen, J. and Vehtari, A. (2016). On the hyperprior choice for the global shrinkage parameter in the horseshoe prior.

Rai, C. S. and Manghnani, M. H. (1981). The effects of saturant salinity and pressure on the electrical resistivity of Hawaiian basalts. *Geophysical Journal International*, 65(2):395–405.

Revil, A., Breton, M. L., Niu, Q., Wallin, E., Haskins, E., and Thomas, D. (2016). Induced polarization of volcanic rocks. 2. Influence of pore size and permeability. *Geophysical Journal International*, 208(2):814–825.

Revil, A., Ghorbani, A., Gailler, L., Gresse, M., Cluzel, N., Panwar, N., and Sharma, R. (2018). Electrical conductivity and induced polarization investigations at kilauea volcano, hawai'i. *Journal of Volcanology and Geothermal Research*, 368:31 – 50.

Simpson, F. and Bahr, K. (2005). *Practical Magnetotellurics*. Cambridge University Press.

Wheelock, B., Constable, S., and Key, K. (2015). The advantages of logarithmically scaled data for electromagnetic inversion. *Geophysical Journal International*, 201(3):1765–1780.